Proc. of the 9th International Symposium on Geotechnical Safety and Risk (ISGSR)
Edited by Zhongqiang Liu, Jian Dai and Kate Robinson
Published by Research Publishing, Singapore.
doi: 10.3850/GRF-25280825_isgsr-090-P258-cd

ISGSR2025

PREDICTINGPORE PRESSURE AT VARYING DEPTHS IN A NORWEGIANRAILWAY PROJECT

Eivind Stein
Geotechnics and Environment, Norwegian Geotechnical Institute (NGI), Oslo, Norway. E-mail: eivind.stein@ngi.no

Per-Anders Hermanrud, Minu Treesa Abraham, Daniel Ryghseter,Thomas Pabst
Geotechnics and Environment, Norwegian Geotechnical Institute (NGI), Norway.

Groundwater monitoring in the form ofpore pressure measurementsis widelyused toassess the impact of construction activities on the surroundings. Pore pressure datacan be used for monitoring groundwater leakage into construction pits and tunnels, as groundwater drawdown might lead to consolidation settlements and building damage depending on the magnitude, duration and spatial extent of the drawdown.

The monitoring process involves installing severalsensors some years beforethe start of groundworkto record seasonal fluctuations and thus establish sensor baselines. Additional piezometers are often installed right before or after the start of construction to cover for missing areas or to replacedamaged sensors.Data from the sensors can be used to generate triangulated/interpolated groundwater surfaces, for analyzing groundwater flow patterns in three dimensions,or for investigating areal extents of groundwater drawdown. However, with a limited number of installed sensors and large areas of interest, pore pressure estimates in areas between observation points are naturally uncertain.

New methods in the realm of machine learning canbe used for pore pressure prediction in such areas.In a Norwegian railway project with 160 installed piezometers, a pore pressure prediction model was tested using a randomforest machine learning algorithm. The model was trained on data from piezometersalong with availablespatial and geological data. Theresults show that the model has moderate but acceptable predictive accuracy with an average R^2 of 0.53 and a mean absolute error of 1.79 mH2O. In addition to being a possible solution forpredictingthe pore pressure at new locations and depths, such model predictions might also be used to validate data from sensors installed late in a project phase.

Keywords: Geotechnical engineering, groundwater, pore pressure, machine learning, prediction model, site investigations.

1. Introduction

Groundwaterdata is important in construction projects for monitoring important groundwater changes. Groundwater changes can be classified intosudden pore pressure increases, or groundwater drawdown caused by leakage into construction pits and underground tunnels.In urban areas, the latter may cause consolidation settlements and damage to infrastructure and buildings(Karlsrud 2002,Myrabø and Moss-Iversen 2014, Langford et. al. 2022).

Pore pressure is commonly measured at different depths using piezometers. One piezometer is usuallyinstalled near the bedrock surface for monitoring the lower aquifer. In some locations, one or more piezometers are also installed higher up in the soil package to measure the vertical pore pressure distribution. The piezometers installed in the lower, often permeable aquifer usually respond quickly to nearby groundwater leakage and are considered the most important for drawdown monitoring purposes.

Common practice is to install some piezometers a few years prior to the beginning of a project to obtain a baseline for monitoring purposes. Additional sensors are usually installed later, some months before or after the start of constructionfor supplementing data and/or for better monitoring coverage in areas of interest. Pore pressure data may be used to generate triangulated or interpolated groundwater surfaces for analyzing groundwater flow patterns or for investigating aerial extents of groundwater drawdown. These methods assume hydrostatic pressure distribution, where pore water pressure increases linearly with depth below the groundwater table, starting from zero pore pressure at the ground water table.Such groundwater surfaces might also be used forgeotechnical designpurposes, however, these simpler methods may beinaccuratedue to topographical effects and the complex nature of groundwater. The methodologies of triangulation and interpolation are intuitive and easy to understand, however, there is an inherent uncertainty related to groundwater estimation based on the observed values at neighboring points. Additionally, the uncertainty increases with increased distance from observation points and may become large in areas with limited data.

The goal of this study has been to investigate the potential for a machine learning model to efficiently predict the pore pressure at unknown points in a local area based on data from a population of installed sensors.

Predicting groundwater levels with machine learning has been done by many authors as summarized by Tao et. al (2022), however, the majority of the studies use time-dependent data to predict groundwater changes with time series models. In this study, the goal hasbeen to predict the average depth to the groundwater table (as pore pressure in mH2O relative to ground surface), on a local scale, using spatial data relationships only.The model was trained on data from piezometers along with available spatial data (terrain elevation, bedrock elevationand soil type thicknesses). The results are compared with an alternative method of radial basis three- dimensional interpolation to assess the potential of the method. The results and recommendations for future studies are discussed.

2. Data

The pore pressure data used in the study is taken from an ongoing Norwegian railway project. A total of 160 piezometers are installed at 70 locations in the X-Y plane. Every location includes a piezometer installed near the lower aquifer above the bedrock surface. The lower aquifer occasionallycontainssand ormoraine depositswitha high permeability. Additionally, 1-3 piezometers may beinstalled at different depths higher up in the soil package, usually in the thicker marine clay layer. Silt and sand layers are also present in the middle clay layer, but to a varying degree. The marine clay layer has generally a low permeability, but the permeability may increasewith the presence of sand and silt horizons. Additionally, the silt content in the clay layer increases with decreasing depth, leading to a higher permeability.A topsoil layer with varying thickness is typically found in the top few meters near the surface. The vertical pore pressure distribution in the area is rarely hydrostatic due to the varying permeability in the deposited layers.

The pore pressure is originally measured in values of kPa, corrected for atmospheric pressure. However, for prediction purposes, the pore pressure is recalculated tometer of water (mH_2O) from ground surface to avoid the obvious dependency on installation depth.The sensors in the study have been installed at different times in the project phase, so the length and timespan of the data series vary. Data outside the timespan of 1^{st}January 2020 - 1^{st} of January 2023 have been filtered out to avoid potential disturbance from construction works. The data have been further resampled toaverage values over the three-year time span. The average value of pore pressure in mH_2O from ground surface is thus the dependent variable to be predicted.

The distribution of pore pressure for the total population of sensors (160) is shown in Fig. 1 (a). As seen in the left figure, the pore pressure distribution variesconsiderably in the project area, ranging from +3 m (above terrain, artesian pressure) to -15 m (below terrain) with a mean of around -2 m.InFig. 1 (b), the pore pressure distribution is plotted against sensor installation depths in 3 m bins. As seen in the figure, the variation in pore pressure is considerable inside each bin and follows no clear pattern in respect to sensor installation depth.

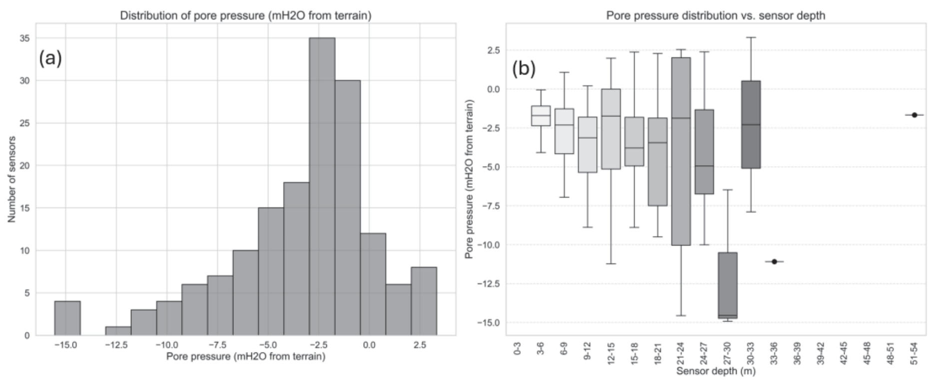

Fig. 1. **(a)**Statistical distribution of pore pressure in mH_2O from ground surface for the entire population of sensors.**(b)** Pore pressure distribution vs. sensor installation depth.

Spatial data is used as a basis for theinput variables in the prediction model. The data is available in the form of raster datasets (GeoTIFF) for the entire project area. The raster datasets include terrain elevation (StatensKartverk 2024) and bedrock elevation, clay thickness, moraine thickness and topsoil thickness (project specific, interpreted from total soundings, CPT and soil samples).A total of 15 data series have been used as external inputs into the machine learningmodel. Some of the datais retrieved directly from the raster datasets at the sensor locations. Other data iscalculated based on theoriginal raster datasets and sensor metadata. Statistical distributions of the inputdata series are shown in Fig. 2.

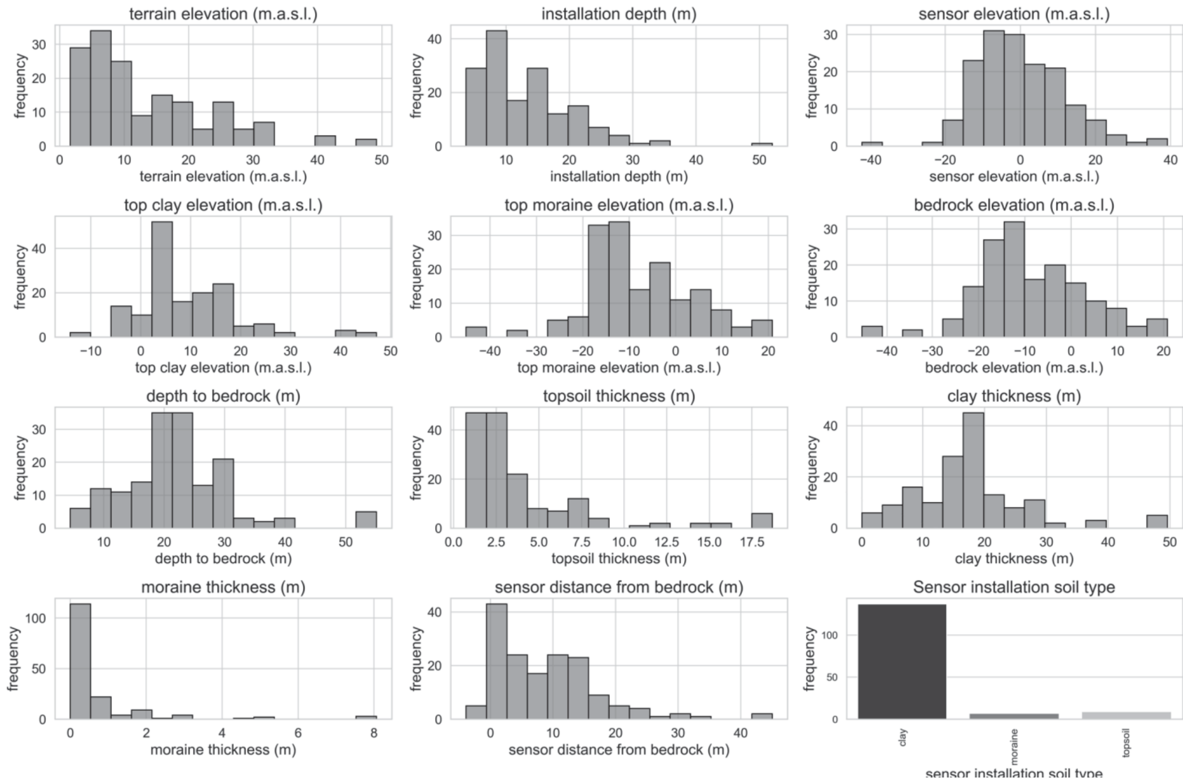

Fig. 2. Statistical distribution of the input variables used as input into the prediction model

3. Methodology

A random forest regression algorithm was chosen for this study. The modelling framework is the open-source python library Scikit-learn with the *RandomForestRegressor*function (Pedregosa et.al. 2011). A total of 40 decision trees make up the random forest model. First, the population of 160 sensors is split into two near-equally large groups (50-50) based on random shuffling of the 70 sensor locations. One group is used for model calibration (training). The remaining group is usedfor validating the model’s ability to predict pore pressure values at unknown locations (testing). The total number of sensors in the train and test group are 75 and 84, respectively.

Prediction accuracy is quantified in terms of the coefficient of determination, R^2, and the mean absolute error, MAE. The R^2 is in simple terms the proportion of explained variability over the total variation(Dodge, 2008). The MAE is simply the average error (residuals) of a model predictions. The MAE is used here as it is considered more intuitive and better suited for assessing model performance than other metrics of error (Willmott and Matsuura2005).

4. Results

The results for a random split of train and test group are shown in Fig. 3.Thepredictive accuracy for the test group is moderate, with a R^2 of 0.68 and a MAE of 1.54 m. It can be seen that the model overpredicts large negative values, but performs better for observations closer to ground water table (0). The residuals (model error) in the test predictions are near normally distributed with a mean close to 0, but with a slight skewed “tail” to the left of the figure (negative error) caused by the obvious overprediction of large negative observations.

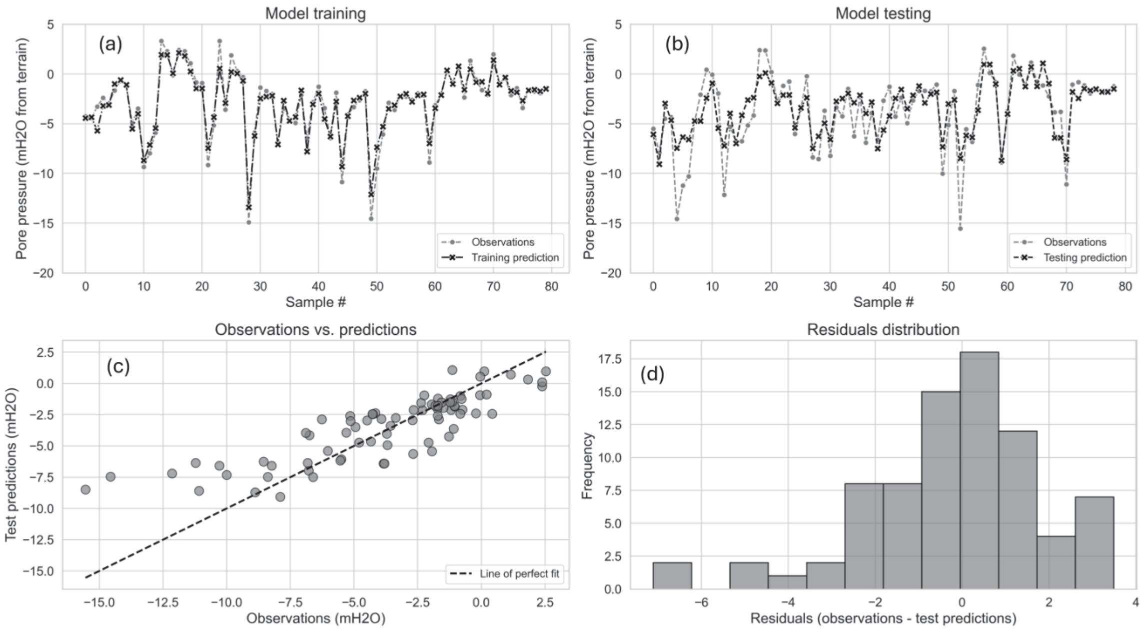

Fig. 3**(a)**Observations vs. prediction for the training group (calibration). **(b)** Observations vs predictions for the test group. **(c)**Scatter of observations vs predictions in the test group and the line of perfect fit. **(d)** Statistical distribution of the model residuals (the model error) for the test group.

The machine learning predictions have been compared with an alternative method of radial basis three-dimensional interpolation at 81 locations (test group) based on 76 known observations. The interpolation method gives a R^2 of 0.35 and an MAE of 2.15 m for the test group.The accuracy of prediction for the two methods is somewhat dependent on the specific shuffling of sensors in the train and test group. When performing 100 different shuffles for the two methods, the random forest model comes out with a mean R^2 of 0.53 and a mean MAE of 1.79. The interpolation method shows an average R^2 of 0.37 and an average MAE of 2.03. The variation in test results is likely a consequence of a small sample size. More data would most likely have reduced the difference in the results between different train-test splits.

4. Discussion

Based on the results, the random forest model outperforms the method of interpolation when looking at overall predictive power for the test set. However, it is likely that the method of interpolation significantly outperforms the machine learning model in areas with a dense population of sensors. Conversely, the machine learning model might be better at predicting values in areas far from known observation points, as the model only considers correlations in the input data series without considering location in the XYZ domain.

The inaccuracy of the machine learning model predictions indicates that the chosenmethodology or model type is inaccurate for describing the system, or that there are other influential variables that have not been input to the model. It is likely a combination of the two. When considering the first point, the amount of data used in the study is considered very small in a machine learning context (but large in a construction project).Aneffect of a small sample size becomes apparent when looking at the varying prediction accuracy of different train and test groups. The predictive accuracy of the model might be significantly improved if the dataset was larger. When considering the latter, the geology in the area is known to be complex.Groundwater levels and flow patterns are also known to be complex and bound with high uncertainty and variability. When considering the key points discussed above, the model prediction accuracy is considered acceptable.The mean error might be too large for the method to be applied directly, however, the prediction accuracy in terms of absolute values might be smallerwhen tested on data from another area withdifferent geological conditions. It is also possible to limit the applicability of the model to depths a range of pore pressures where the residuals are between a specific range (For example, (-1,1)). It should be noted that the data used for training should be site specific, and large datasets are required to develop generalized models which can be used for multiple geological settings. A machine learning model such as the one used in this study should always be used with caution, as a model may pick up on correlations that might be caused bycoincidence rather than causation. This is especially true when having small datasets.However, the authors see potential for the methodology to be tested further, in new areas and in new ways.

Acknowledgement
This work was funded by the basic funding to NGI from The Research Council of Norway.

References

Dodge, Y. (2008). Coefficient of Determination. In: The Concise Encyclopedia of Statistics. Springer, New York, NY. https://doi.org/10.1007/978-0-387-32833-1_62

Karlsrud, K. (2002) Control of water leakage when tunnelling underurban areas in the Oslo region. Water Control in Norwegian Tunnelling,Publication No. 12. Norwegian Tunnelling Soc 27–33

Langford, J. Holmøy, K.H., Hansen, T.F., Holter, K.G., Stein, E. (2022) Analysis of water ingress, grouting effort, and pore pressurereductioncaused by hard rock tunnels in the Oslo region. Tunneling and Underground SpaceTechnology 130:104762

Myrabø, S, Moss-Iversen, E. (2014) The Romeriksporten tunnel. Settlementsas a result of tunnel construction and currentstatus.Norwegian Tunnelling Society. In Proceedings of 2014 Fjellsprengningskonferansen 20.1–20.10. Oslo, Norway.

Pedregosa, F., Varoquaux, Ga"el., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. Scikit-learn: Machine learning in Python. Journal of machine learning research. 2011;12(Oct):2825–30.

Statenskartverk (2024). Public terrain elevation data for Norway. Accessible online from https://hoydedata.no/

Tao, H., Hameed, M.M., Marhoon, H.A., Zounemat-Kermani, M., Heddam, S., Kim, S., Sulaiman, S.Q., Tan, M.L., Sa'adi, Z., Mehr, A.D., Allawi, M.F., Abba, S.I., Zain, J.M., Falah, M.W., Jamei, M., Bokde, N.D., Bayatvarkeshi, M., Al-Mukhtar, M., Bhagat, S.K., Tiyasha, T. and Yaseen, Z.M. (2022). Groundwater level prediction using machine learning models: A comprehensive review., *Neurocomputing*, 489, pp. 271–308. Available at: https://doi.org/10.1016/j.neucom.2022.03.014.

Willmott, C.J. and Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research Vol. 30*, 79-82.